Transformer-based Joint-station Seismological Analysis Applied to Focal-mechanism Determination

Xiaohan Song[1], Men-Andrin Meier[2], William L. Ellsworth[1], Gregory C. Beroza[1]

[1] Department of Geophysics, Stanford University, CA, 94305 [2] Institut für Geophysik, ETH Zürich, 8092 Zürich, Switzerland

1. Motivation

Deep learning models have been widely applied in processing seismic signals. But when it comes to combining phase information (e.g., arrivals, polarities, amplitudes) across a station network, most workflows rely on either physics-informed optimization (e.g., grid search), or neural networks trained on fixed station arrays (which limits their application to specific areas).

We propose that by treating every station's phase information as a "token vector" and appending them with the station locations/metadata as "positional encodings", we can train a transformer-based neural network capable of processing phase information across any distributed station arrays.

We aim to build a transformer-based focal mechanism (FM) determination model and validate it on an FM dataset from the 2016 Italian Amatrice sequence [1].

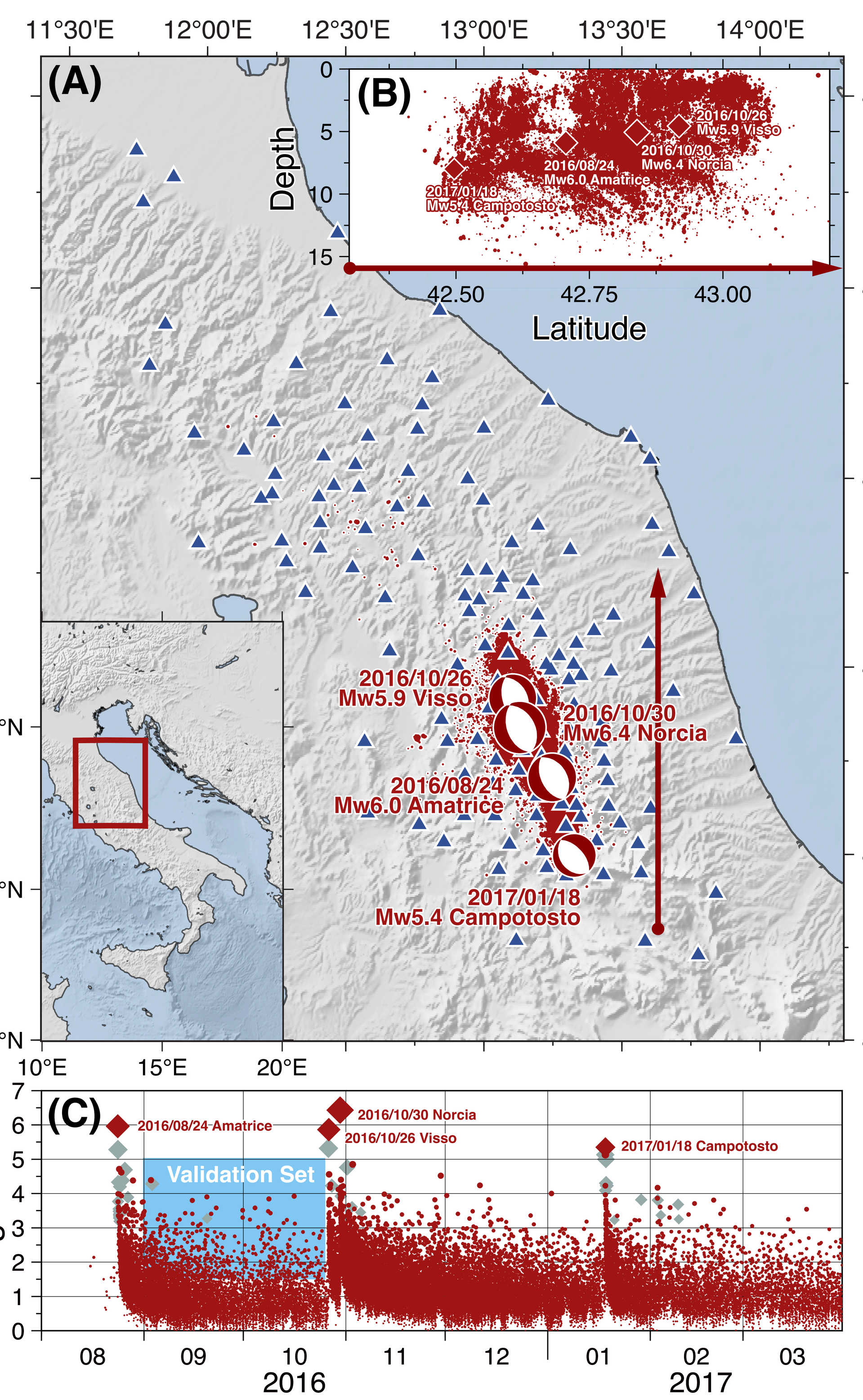

Figure 1: (A): The Amatrice sequence seismicity and the local INGV stations in Meier et al.'s FM catalog [1]. (B): Cross-sectional view. (C): Temporal distribution of the sequence.

References

[1] Meier et al., A deep catalogue of 56k focal mechanisms for the. . . . *EGU*, 2023.

[2] Xi et al., PyFK: A Fast MPI and CUDA Accelerated. . . . *AGU*, 2021.

[3] Reasenberg and Oppenheimer, FPFIT, FPPLOT and FPPA. . . . *U.S. Geological Survey*, reports 1985.

[4] Hardebeck and Shearer, Using S/P Amplitude Ratios to. . . . *BSSA*, 93(6), 2434–2444, 2003.

[5] Skoumal et al., SKHASH: A Python Package for Computing. . . . *SRL*, 95(4), 2519-2526, 2024.

2. FocoNet Model Structure

Our transformer-based model, named FocoNet, whose key part consists of 7 layers of multihead self-attention, directly solves for the focal mechanism based on the relative locations (to the source), first-motion polarities, S/P amplitude ratios, and SNRs from a set of stations. Ideally, we assume that the polarities have been determined by a pre-existing neural network (e.g., Meier et al.'s CNNs [1]), and we know the source locations & phase arrivals.

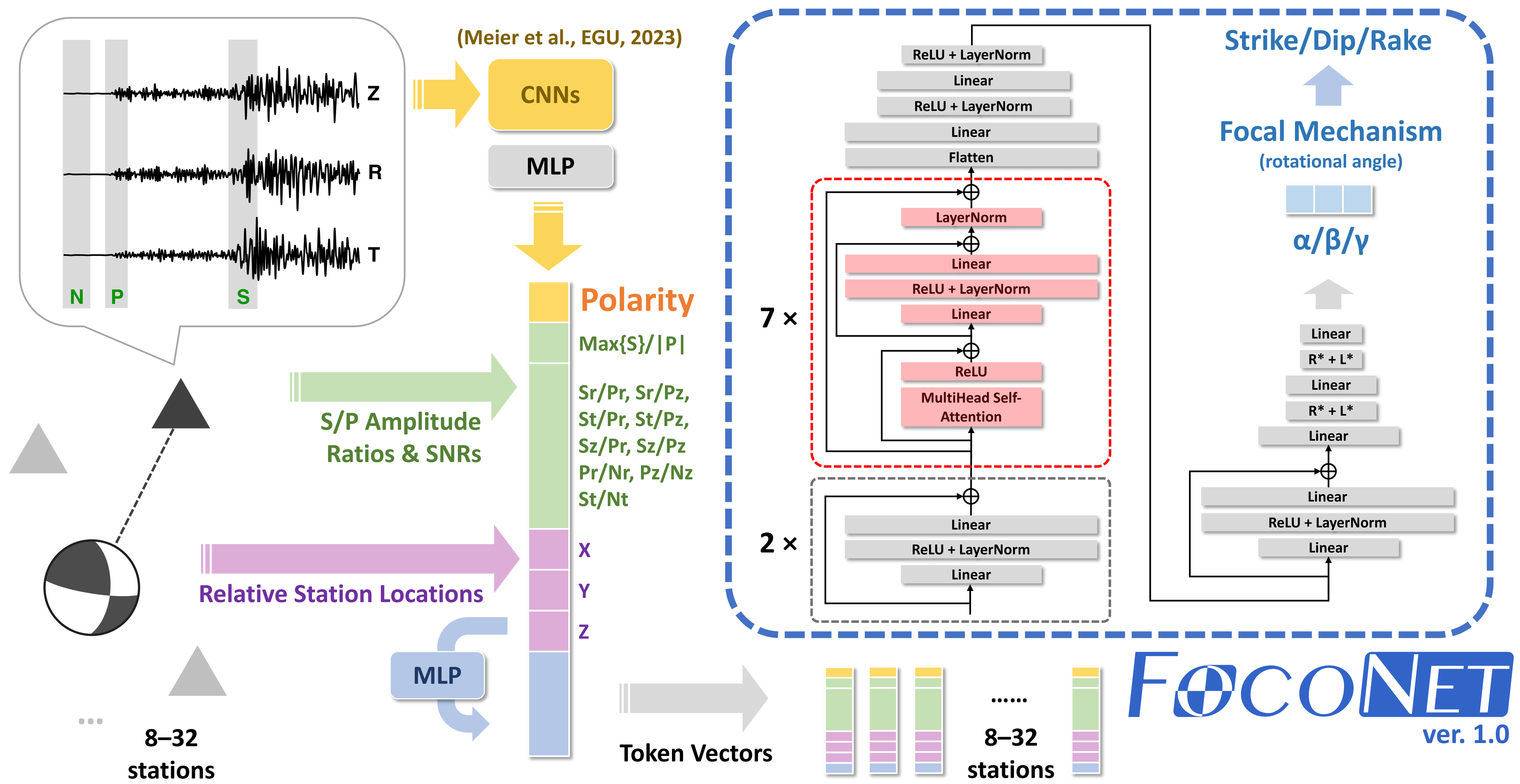

Figure 2: The workflow to solve for a focal mechanism and the model structure of FocoNet

3. Training and Test

To train FocoNet, we create 500,000 randomly rotated focal mechanisms. For each event, we randomly distribute 8–32 stations within 120-km epicenter distance. We generate synthetic waveforms through PyFK [2]. We randomly added noise to the synthetics to mimic the Amatrice real waveform as our training and test sets. We use the average (L1 norm) Kagan angle between our predictions and known labels as the Loss function.

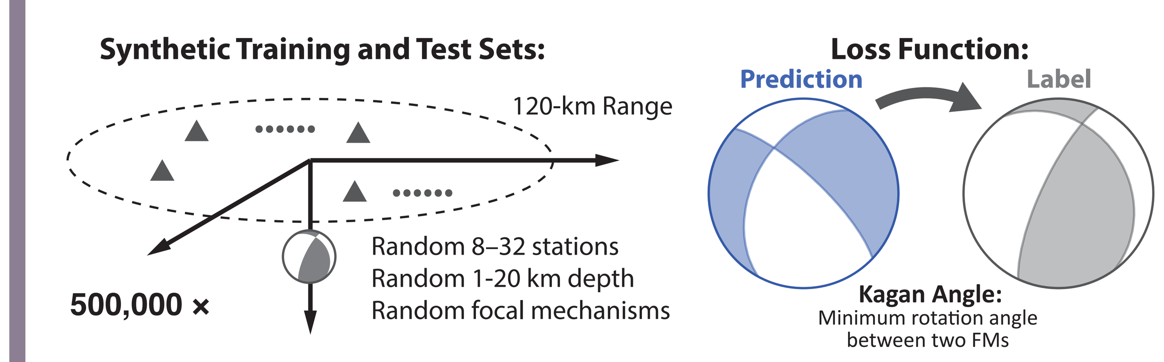

Figure 3: Our training/test set and loss function

We test FocoNet on different synthetic test cases and compare the predictions with those from traditional methods (e.g., FPFIT [3] and (SK)HASH [4, 5]). FocoNet consistently gives the most accurate focal mechanisms in all of the test cases.

Kagan Angle (°) to known label FMs (Mean / Median)

Test Sets 512 × / Methods	FPFIT	SKHASH (excluding S/P)	SKHASH (including S/P)	FocoNet
12 stations, 50 km	32.38 / *26.01*	35.91 / *30.82*	29.43 / *23.09*	**18.08** / ***14.16***
24 stations	21.79 / *15.87*	20.25 / *16.44*	17.30 / *13.26*	**14.45** / ***9.28***
32 stations	18.67 / *13.99*	17.11 / *12.76*	14.45 / *11.70*	**9.98** / ***8.59***
24 stations, 90° Azimuth Gap	24.09 / *17.83*	23.39 / *18.29*	19.28 / *14.41*	**14.00** / ***11.45***
24 stations, 180° Azimuth Gap	29.76 / *24.06*	28.14 / *23.34*	22.72 / *17.12*	**17.40** / ***14.17***
24 stations, 270° Azimuth Gap	43.97 / *40.46*	43.50 / *42.05*	34.29 / *27.74*	**30.43** / ***24.62***
12 stations, 10% Polarities Flipped	40.61 / *33.98*	40.89 / *35.24*	33.25 / *25.06*	**23.04** / ***17.66***
24 stations, 10% Polarities Flipped	30.07 / *23.16*	25.77 / *20.02*	21.73 / *15.77*	**14.60** / ***12.03***
32 stations, 10% Polarities Flipped	26.06 / *21.02*	22.12 / *16.50*	18.08 / *13.01*	**13.08** / ***10.31***

Table 1: Comparing different methods from the Kagan angles between predicted and (known) label FMs

4. Validation on the Amatrice Sequence

We validate FocoNet on 300+ Amatrice sequence events with A-class HASH solutions and 600+ events with D-class solutions. The Kagan angle between the FocoNet predictions and the corresponding A-class HASH FMs is on average 18°. For HASH FMs with quality class D, the distribution of Kagan angles to the reference FM (325/39/-97) is relatively uniform, while FocoNet predictions have a Kagan angle distribution that peaks at 30°.

The transformer-based FocoNet can efficiently combine different kinds of phase information across a randomly distributed station array, and give more accurate focal mechanism solutions than traditional grid search methods can do, even with noisy data.

Both the enriched types of phase information (HASH standard S/P ratios, S/P ratios from different channels, and SNRs) and the increased ability to address uncertainty make FocoNet a promising alternative to tradition methods.

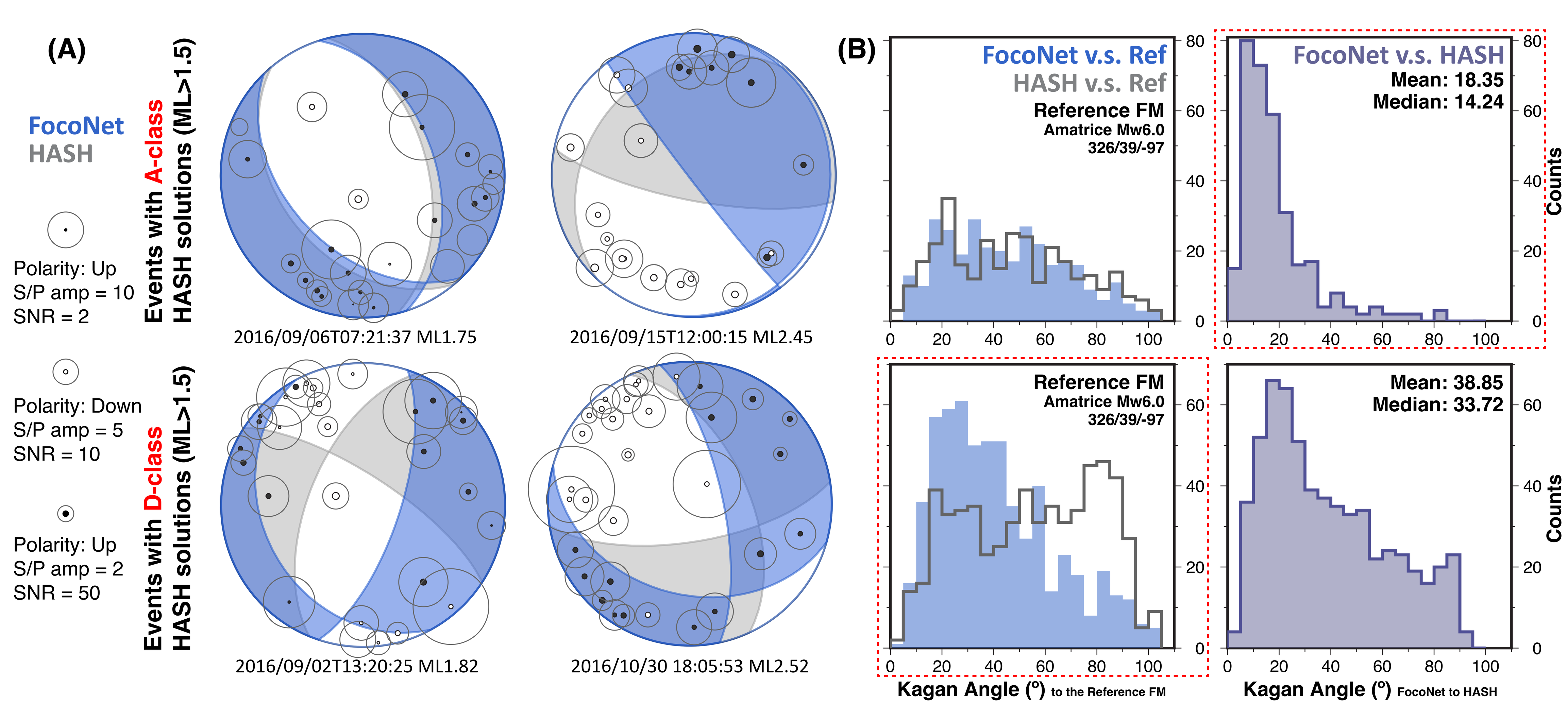

Figure 4: (A): Examples showing the comparison between FocoNet and HASH solutions. (B): Histograms of the Kagan angles between FocoNet/HASH solutions and the reference focal mechanism, and between FocoNet and HASH solutions.

Contact

Xiaohan Song: xhsong@stanford.edu